

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-296311

(43) 公開日 平成11年(1999)10月29日

(51) Int.Cl.⁶

G 0 6 F 3/06

識別記号

3 0 4

F I

G 0 6 F 3/06

3 0 4 B

審査請求 未請求 請求項の数 3 O L (全 5 頁)

(21) 出願番号 特願平10-95689

(22) 出願日 平成10年(1998)4月8日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(71) 出願人 000233055

日立ソフトウェアエンジニアリング株式会
社

神奈川県横浜市中区尾上町6丁目81番地

(72) 発明者 藤本 健雄

神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内

(74) 代理人 弁理士 小川 勝男

最終頁に続く

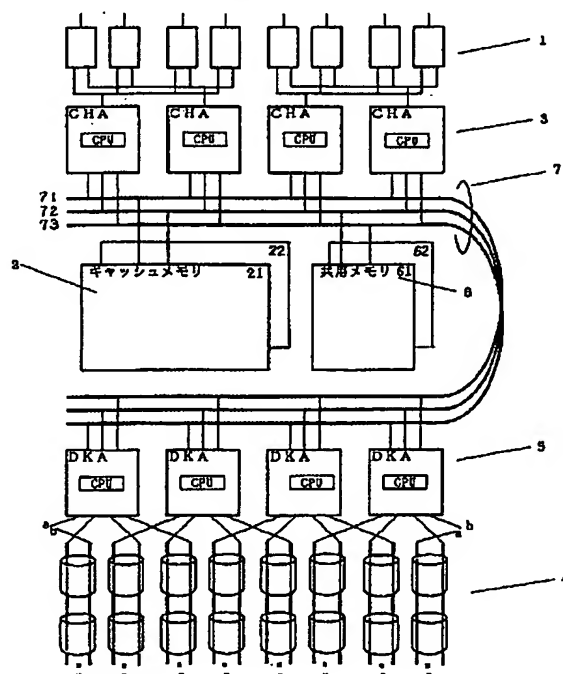
(54) 【発明の名称】 記憶装置の耐故障制御方式

(57) 【要約】

【課題】 記憶装置における耐障害制御方式の一手法を提案し、複数の論理バスがあり、共通部位を含む記憶装置における障害発生に対し、正しい故障部位の特定、切り離しをオンライン動作中に自動的に行う。それにより、障害の続発を抑止し、システムの信頼性を向上させる。

【解決手段】 システムを構成し、切り離し可能な各部位毎に障害検出回数を記憶するカウンタを設け、またそのカウンタ値を総合的に判定する論理を具備する。本発明方式は部位毎の障害発生回数の統計より、故障部位の高精度の特定を目的とし、それにより故障部位をシステムより排除し、正常部位を用いたシステム動作を継続可能とする。

図 1



【特許請求の範囲】

【請求項 1】バス結合により、複数コンポーネントからなる記憶装置制御部において、障害発生回数を切り離し可能な部位毎に計数し、またその計数結果を用いた統計的分析により、故障発生部位の判定およびシステムからの切り離しを自動的に行う耐故障制御方式。

【請求項 2】上記請求項 1 の故障部位判定方式において、障害部位切り離し後の障害発生状況を引き続き監視し、故障部位判定結果が不適切と認識した場合には、不当に切り離した部位の自動回復を行うなど、継続的なフィードバックを可能とする耐故障制御方式。

【請求項 3】1 回の障害検出のみによる故障部位特定が困難な場合において、上記請求項 1、請求項 2 の制御方式を用い、故障部位の誤判定を防ぎ、システムダウンを回避可能なことを目的とする記憶装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】その一方、コンピュータシステムの構成において、複数論理アクセス経路（以下では論理バス）を有しながらも物理的に共通部位を持つことが多くなっている。例えば、構成各コンポーネントを共通バスに接続するバス結合方式は、システムの構成変更、拡張が容易のため広く用いられている。しかし共通部位を持つことにより、個別の障害が全体に影響を及ぼしたり、1つの障害発生が他の障害を併発させたりするなど、障害部位の特定も難しくなっている。

【0002】本発明はこのような障害部位特定の問題に対し、1つの解決方法を提案している。障害部位特定の精度を向上させることにより、故障部位の切り離しが正確に行え、結果的にシステム全体の信頼性をも向上可能となる。

【0003】

【従来の技術】近年、記憶装置に対する信頼性の要求が高まっている。特に無停止システムの構築に際し、高信頼性を持つ記憶装置は欠かせないものとなった。そのため制御部を含むシステムの各構成要素はすべて冗長性を持たせ、故障部位さえ特定し、切り離せれば、代替コンポーネントによるシステム動作が継続可能となる。

【0004】検出した障害内容により、1回の障害検出ですぐ故障部位を特定し、当該部位を閉塞する方式が多く用いられている。しかし、そのためには複雑で高度な障害検出構成が必要となり、また恒久的に発生しない一時的なノイズにもすぐに閉塞を行い、システムの冗長度を下げかねない。

【0005】それに対し、ハード的な部位ごとに障害検出回数を累計し、あらかじめ設定された固定しきい値を越えれば、当該部位を故障発生部位と判定してシステムより切り離す方式もあった。しきい値を越えない回数、一時的なノイズでは閉塞に結びつかないので過閉塞をある程度回避可能である。但し、システム共通部位を

有し、各部位が複雑に絡み合うシステムにおいては、1カ所の障害で複数の論理バスに影響を来すことがあり、複数の部位で障害回数が加算されることも考えられる。その場合、単純なしきい値チェックのみで故障部位を正確に特定することは難しい。

【0006】

【発明が解決しようとする課題】記憶装置における障害リカバリ処理で特に重要なのは、システム動作しながら故障部位を抽出、切り離すことである。しかし、システムを構成する各コンポーネントを共通バスに接続するバス結合方式など、共通部分を持つシステムにおいては、単体の故障によってシステム全体に障害が波及することがあり、故障していない部位も一時的に正常動作できない恐れがある。その場合においても、障害元となる故障部位をいかに正しく判定し、障害の再発を断ち切るかが、システム信頼性確保における重要な課題である。

【0007】複数論理バスがあるシステムにおける単体故障発生時、本発明方式の適用により、故障部位を高い確率で指摘可能となり、当該故障部位をシステムから切り離すことで、システムの継続動作を保証する。また、共通バスなど特に各論理バスとも複雑に絡み合う部位については、故障部位の閉塞後にも引き続きに障害発生状況を監視でき、障害が収束しない時には一旦閉塞したバスを回復するなどリカバリを行う。

【0008】

【課題を解決するための手段】上記課題の達成を目的とし、本発明ではシステムを切り離し可能な複数部位に分け、それぞれの障害発生回数を計数し、障害回数の総合的な判断から故障部位の特定を行う。ここで総合的な判断というのは、1つの部位の故障判定についてはその部位の障害発生回数のみならず、他の部位およびシステム全体の障害発生回数とも比較、判定するため、バスなど共通部位が故障した場合は、それを使用する各論理バスで障害が検出されることが予想され、単一部位の障害検出回数が先にしきい値を越えても他の部位と比べて回数突出して大きくなれば、共通部位故障の疑いがあり、簡単に当該単一部位を閉塞しないような論理とする。

【0009】具体的には、例えばある共通部位を共用する複数バスがあり、それぞれの使用頻度がほぼ同一レベルであるとする。共通部位の故障時、各バスにおける障害発生回数は全体の算数平均に近いはず、1つのバスにのみ影響する単一部位の障害時、当該バスの障害発生回数だけが大きく、他のバスの障害回数がほとんどないはず。実際各バスに計数された障害回数を、前記期待結果との相似を計算すれば、共通部位の障害が、当該単一部位の故障かを判定可能である。

【0010】複数論理バスを備える記憶制御装置におき、本発明方式を用いることにより、複数バス共通部位の故障か、単一バスのみに影響する個別部位の故障かを区別することが可能。個別部位の故障であれば、当該論

理バスを閉塞してシステムから切り離すなどを行い、システム全体への影響を抑止する。共通部位の故障と判定したとき、冗長度を具備したシステムであれば、その共通部位の部分縮退を施すことにより、不要なバス閉塞を行うことなくシステム動作を続行できると考えられる。

【0011】また本発明方式により、共通部位の部分縮退を実行した後も障害発生回数を統計し、障害の発生が収束しない場合には縮退した共通部位を回復し、共通部位の他の部分または個別部位の閉塞を行うことが可能。常に同時に使用され、障害発生回数より切り分けできない複数共通部位が存在する場合でも、先ず1つの部位を閉塞させてから、その後の障害発生を監視すれば、正しく障害部位を排除できたかどうかを検証可能である。

【0012】

【発明の実施の形態】本発明方式の実施例について、図面を用いて詳細に説明する。

【0013】図1は、本発明方式を適応した記憶制御装置のブロック図である。ホストコンピュータに接続するチャンネル接続系1、入出力データを一時的に格納するキャッシュメモリ2、チャンネル接続系1とキャッシュメモリ2間のデータ転送を制御するチャンネルアダプタ（以下ではCHA）3、データを蓄積する記憶媒体であるディスクアレイ4、キャッシュメモリ2とディスクアレイ4間のデータ転送を制御するディスクアダプタ（以下ではDKA）5、システム管理情報および通信情報などを格納する共用メモリ6、各CHA3、DKA5からキャッシュメモリ2または共用メモリ6へ接続する共通バス7によって構成される。

【0014】単一部位の故障によるシステム動作停止を防ぐため、各構成要素は冗長性を持たせている。すなわち、CHA3とDKA5はシステム内にそれぞれ複数存在する。

【0015】キャッシュメモリ2と共用メモリ6はそれぞれ切り離し可能な2面化構成となっている。ここではキャッシュメモリA面21、キャッシュメモリB面22、共用メモリA面61、共用メモリB面62と呼び、どの片面故障時にも正常の1面のみで動作可能である。

【0016】ディスクアレイ4はパリティディスクを含み、任意の1つのディスクを閉塞させて継続運転できる。

【0017】共通バス7はHバス71、Lバス72、Mバス73の3本のバスから構成される。Hバス71はキャッシュメモリ2アクセスに用いられ、Mバス73は共用メモリ6アクセス用に用いられる。通常、Lバス72は予め設定されているシステムオプション情報の指定により、Hバス71と協調してキャッシュメモリ2の高速アクセス（同時使用によって2倍のバス幅を実現）に、または独立して共用メモリ6アクセスに使用可能である。Hバス71が故障時、Lバス72をキャッシュメモ

リ2アクセスに用い、Mバス73が故障時共用メモリ6アクセスに用いることで、1本のバスが故障してもシステムが継続動作できる。

【0018】本実施例システム稼動時には、複数のアクセス論理経路（以下では論理バス）を用いた多くのアクセスが同時に動作する。例えば、チャンネル接続系1よりデータのあるCHA3を経由してキャッシュメモリA面21へ転送する論理バス、他のCHA3から共用メモリB面62へ転送する論理バス、またはキャッシュメモリB面22からあるDKA5を経由してディスクアレイ4へ転送するバスなど、様々なバスを用いたアクセスが同時に動作する。

【0019】システム動作中、ある特定ハード部位に故障が発生し、当該部位を用いた論理バスが恒久的にアクセス不可となれば、障害切り分けテストによって、故障部位を特定してシステムから切り離すことができる。ここでいう障害切り分けテストとは、例えば障害検出バス上の1つの部位だけを他のものに切り替えてアクセス試行し、その結果から当該部位の故障かどうかを判定する論理である。一例として、あるCHA3からHバス71とLバス72を同時に用いるバスモードでキャッシュメモリ2へデータ転送を行う際に、データパリティエラーの障害を検出すると、障害切り分け処理において、当該CHA3からHバス71とLバス72を経由し、それぞれキャッシュA面21とキャッシュB面22へアクセステストを行う。例えば、Hバス71を用いた2面のアクセスに障害がともに検出され、Lバス72を用いた2面のアクセスに障害が検出されなければ、Hバス71の故障と判定できる。その場合、Hバス71を閉塞し、Lバス72のみをデータ転送用に切り替えて、システムは継続動作可能となる。

【0020】但し、実際のハード故障は必ずしも恒久的なものではなく、一時障害が多発することがある。また特定タイミング、特定アクセスパターンのみ障害が発生することも考えられる。その際、前記障害切り分けテストで故障部位を使用したアクセステストにおいても正常終了し、故障部位を特定できない場合が多い。

【0021】ここで、障害切り分けテストより部位特定が可能なもの以外、一時障害などのケースに対して、障害発生回数をカウントし、総合的なしきい値判定を用いて故障部位の特定を行う。

【0022】以下はまず障害発生回数の計数方式を述べ、次に総合しきい値判定の論理を説明する。

【0023】説明の便宜上、本実施例記憶制御装置における障害を、データ転送系障害と共用メモリ情報アクセス障害の2種類に限定する。各部位における障害発生回数を計数するため、各障害種別につき、システムを構成する部位種別毎に1種類のカウンタを割り当て、また各種別内に存在する要素数毎に障害回数を別々に累積する。例えば、あるCHA3における共用メモリ情報アク

セス障害の発生回数、キャッシュメモリA面21におけるデータ転送系障害の発生回数、Lバス72におけるデータ転送障害の発生回数など、それぞれカウントする。本方式の特徴として、単にシステム全体の障害回数を累積するだけでなく、障害検出時に使われている論理バスを解析し、当該経路上に存在する各部位のカウントアップを行う。

【0024】いま、例えばある特定CHA3のデータ転送系機能が故障し、一時障害が多発するケースを考える。そのとき、当該CHA3のデータ転送障害カウンタのみが高い値を示し、他のCHA3またはDKA5のカウンタはカウントアップされない。なお、キャッシュメモリ2の両面は通常ほぼ均等な確率にアクセスされるため、故障CHA3からのデータ転送が両面に分散され、キャッシュメモリA面21とキャッシュメモリB面22における障害発生カウンタはほぼ同じ値を取るものと考えられる。

【0025】また、例えばHバス71とLバス72をデータ転送に用い、Mバス73を共用メモリ情報アクセスに規定するシステムに、Mバスに故障が生じたケースを考える。この場合、Mバス73は完全なシステム共通部位であり、動作中の全CHA3およびDKA5はMバス73を経由して共用メモリ6へのアクセスを行うため、各CHA3、DKA5で障害が均等に検出され、共用メモリA面61と共用メモリB面62における障害回数もほぼ同じ値になる。

【0026】障害回数カウンタは上記2種類の傾向があることを踏まえ、システム内各カウンタ値を用い、故障部位を特定する判定論理を説明する。

【0027】まず、各部位の障害回数についてベースとなるしきい値を設ける。当該障害回数カウンタ値がそのしきい値に達しない場合は、正常システムにおいても発生し得るノイズの可能性もあるとし、ある部位の障害カウンタがベースしきい値を越えた時点で総合判定を行う。

【0028】判定に際しては、ベースしきい値を越えたカウンタと同一種類のカウンタをすべて抽出して、判定の材料に用いる。ここで同一種類のカウンタは同等機能を持つ、独立に動作するその他の部位のカウンタを指す。例えば、あるCHA3のデータ転送障害回数がベースしきい値を超過した場合、他の動作中のCHA3のデータ転送障害カウンタ値をも取り出す。それらを元に最初に障害検出した部位のカウント値が他のカウント値に比べて飛び抜けて大きくなっているか、それとも各カウンタの値が近い値を示しているかを判定する。

【0029】その判定方法の一例として、現実の障害回数分布が上記2ケースの理想回数分布にどれだけ近いかを、相似度の計算で比較する方法がある。個別部位故障のケースでは、理想的には当該部位の障害検出回数のみが抽出したカウンタの合計値に達し、他のカウンタ値が

0となる。ここで、実際の各部位のカウント値と上記理想値の距離（差値の2乗）の合計を求め、この計算値が小さいほど個別部位故障の理想分布に近い。あらかじめ設けられる基準値以上に、上記理想分布に近ければ、当該部位の故障と判定する。同様に、共通部位故障のケースでは、理想的には全カウンタ値が各部位の平均回数に等しい。同じように実際のカウント値から上記理想値の距離（差値の2乗）の合計を求めれば、共通部位故障の理想分布との相似度が得られる。基準値以上に近ければ、共通部位の故障と判定する。

【0030】ここで本実施例では、上記2種類の理想分布との比較を行っているが、他の障害分布を持つハード構成も考えられる。例えば共通部位の故障において、特定ある部位は他部位の2倍の確率に障害が検出される場合なども、対応した理想分布を用意すれば、同様な相似度計算が可能となる。さらに、実際各部位へのアクセス数を計数し、そのアクセス回数に応じた障害予想回数を動的に、前記理想分布に反映する方式も容易に考えられる。

【0031】以上の判定を経て、個別部位の故障と判定すれば、当該部位をシステムから切り離す。例えばキャッシュメモリA面21の故障と特定したら、キャッシュメモリA面21を使用しないようにシステム管理情報を更新し、正常なキャッシュメモリB面22のみによるシステム動作を行う。

【0032】共通部位故障と判定したときにも、可能な限りその共通部位を縮退して全体に影響を与えないようにシステム動作を継続する。例えば、共用メモリ情報アクセス障害からMバス73の故障と判定したとき、Lバス72を共用メモリアクセス用に切り替え、Mバス73を介したアクセスを停止する。

【0033】故障部位を判定し、障害閉塞を行った後には障害発生回数のカウンタ値をクリアする。

【0034】なお、共通部位が同時に使われ、故障部位が一意的に決定できないケースも考えられる。例えば、本実施例システムは冒頭に言及したように、Hバス71とLバス72を合わせた高速転送バスモードが指定可能である。そうしたシステムオプションが指定されたときに、Hバス71もしくはLバス72の単体の故障でも、2バスを同時に使うために2バスに同じ障害回数が計数されていて、どのバスの故障かを切り分けることができない。このケースを考慮し、Hバス71を先ず閉塞させる論理を盛り込む。そして当該閉塞実行後の障害発生状況を引き続きに監視し、もう一度ベースしきい値を越えて同じように共通部故障と判定したときに、前回の判定結果を引き継ぎ、Hバス71を回復して、Lバス72を閉塞するように制御する。

【0035】このような、閉塞後の障害発生状況を監視し、一旦閉塞された部位を再び回復する論理は、前記同時使用された複数部位の切り分けに必要なのみではなく、

何らかの要因によって誤った故障部位指摘後の訂正にも有効なのは明らかである。

【0036】

【発明の効果】本発明方式を適用した記憶制御装置は、システムで発生する障害を各部位毎に回数の統計を行い、それを用いて故障部位の特定を含む一連の動作を、システム稼動中に自動的に行うことが可能である。これにより、共通バス接続など共通部位を持ち、故障箇所の特定が困難なシステムにおいても、故障部位を高精度で判定でき、その結果故障部位をシステムから切り離し、障害の続発は抑止可能となる。

【図面の簡単な説明】

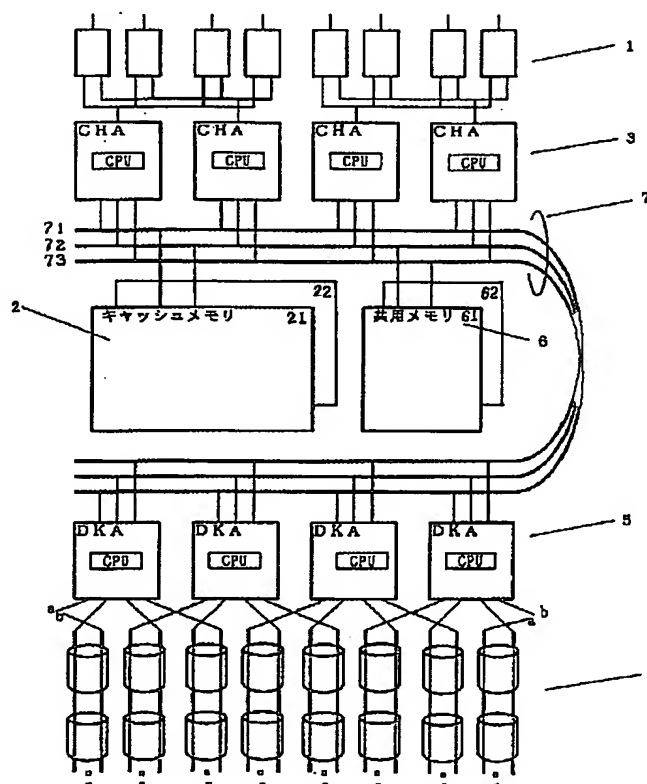
【図1】実施例記憶制御装置のブロック図。

【符号の説明】

1…チャネル接続系、 2…キャッシュメモリ、 3…CHA（チャネルアダプタ）、 4…ディスクアレイ、 5…DKA（ディスクアダプタ）、 6…共用メモリ、 7…共通バス、 21…キャッシュメモリA面、 22…キャッシュメモリB面、 61…共用メモリA面、 62…共用メモリB面、 71…Hバス（データ転送専用バス）、 72…Lバス（データ転送、共用メモリアクセス用切り替え可能バス）、 73…Mバス（共用メモリアクセス専用バス）。

【図1】

図1



フロントページの続き

(72)発明者 本間 久雄
神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内

(72)発明者 阪口 治
神奈川県横浜市中区尾上町6丁目81番地
日立ソフトウェアエンジニアリング株式会
社